

A meta-analytic review of self-reported, clinician-rated, and performance-based motivation measures in schizophrenia: Are we measuring the same “stuff”?

Lauren Luther<sup>a\*</sup>, Ruth L. Firmin<sup>a</sup>, Paul H. Lysaker<sup>b, c</sup>, Kyle S. Minor<sup>a</sup>, & Michelle P. Salyers<sup>a</sup>

\*Corresponding author. Contact information: 402 N. Blackford St., Indianapolis, IN 46202 United States of America. Email: [lutherl@iupui.edu](mailto:lutherl@iupui.edu). Telephone: 317-278-4611. Fax: 317-274-6756

<sup>a</sup> Indiana University-Purdue University at Indianapolis (IUPUI), 402 N. Blackford St., Indianapolis, IN 46202 United States of America

<sup>b</sup> Indiana University School of Medicine, 340 W 10th St #6200, Indianapolis, IN 46202, United States of America

<sup>c</sup> Roudebush VA Medical Center, 1481 W 10th St, Indianapolis, IN 46202, United States of America

---

This is the author's manuscript of the article published in final edited form as:

Luther, L., Firmin, R. L., Lysaker, P. H., Minor, K. S., & Salyers, M. P. (n.d.). A meta-analytic review of self-reported, clinician-rated, and performance-based motivation measures in schizophrenia: Are we measuring the same “stuff”? *Clinical Psychology Review*. <https://doi.org/10.1016/j.cpr.2018.04.001>

### Abstract

An array of self-reported, clinician-rated, and performance-based measures has been used to assess motivation in schizophrenia; however, the convergent validity evidence for these motivation assessment methods is mixed. The current study is a series of meta-analyses that summarized the relationship between methods of motivation measurement in 45 studies of people with schizophrenia. The overall mean effect size between self-reported and clinician-rated motivation measures ( $r = .27, k = 33$ ) was significant, positive, and approaching medium in magnitude, and the overall effect size between performance-based and clinician-rated motivation measures ( $r = .21, k = 11$ ) was positive, significant, and small in magnitude. The overall mean effect size between self-reported and performance-based motivation measures was negligible and non-significant ( $r = -.001, k = 2$ ), but this meta-analysis was underpowered. Findings suggest modest convergent validity between clinician-rated and both self-reported and performance-based motivation measures, but additional work is needed to clarify the convergent validity between self-reported and performance-based measures. Further, there is likely more variability than similarity in the underlying construct that is being assessed across the three methods, particularly between the performance-based and other motivation measurement types. These motivation assessment methods should not be used interchangeably, and measures should be more precisely described as the specific motivational construct or domain they are capturing.

*Keywords: apathy; avolition; psychosis; meta-analysis; measurement*

Among people with schizophrenia, motivation deficits have long been observed (Bleuler, 1911/1950; Kraepelin, 1913/1919) and are among the strongest determinants of reduced functioning and quality of life (Fervaha, Foussias, Agid, & Remington, 2014; Foussias et al., 2011; Kirkpatrick, Fenton, Carpenter, & Marder, 2006). The functional significance of motivation, coupled with the consensus statement on negative symptoms from the National Institute of Mental Health and Treatment Research to Improve Cognition in Schizophrenia (Kirkpatrick, Fenton, Carpenter, & Marder, 2006), has contributed to a burgeoning area of research focused on the assessment of motivation in schizophrenia. To date, a panoply of measures has been used to assess motivation in schizophrenia, including self-reported, clinician-rated, and performance-based measures. However, the degree to which these measures assess a similar underlying construct is not clear (J. Choi, Choi, Reddy, & Fiszdon, 2014; Fervaha, Foussias, Takeuchi, Agid, & Remington, 2015; Kremen, Fiszdon, Kurtz, Silverstein, & Choi, 2016). Determining the overlap or discrepancy between the underlying construct of these measures is needed to clarify the nature of motivation deficits in schizophrenia and develop effective treatments for these deficits. If measures are indeed assessing disparate constructs or constructs that are not directly related to motivation, then using them to evaluate treatments may provide little insight into the effectiveness of an intervention. Thus, these meta-analyses examine the degree of convergence between self-reported, clinician-rated, and performance-based measures of motivation used in existing studies involving people with schizophrenia-spectrum disorders.

### **Conceptualization of Motivation in Schizophrenia Research**

Although numerous definitions of motivation have been proposed, motivation is generally defined as an internal state that initiates, directs, and maintains goal-directed behavior (Kleinginna & Kleinginna, 1981). Within schizophrenia research, motivation deficits are often described as falling under the umbrella of negative symptoms but have been found to be partially independent of negative symptoms (Luther, Lysaker, Firmin, Breier, & Vohs, 2015; Saperstein, Fiszdon, & Bell, 2011; Yamada, Lee, Dinh, Barrio, & Brekke, 2010). Recently, researchers have begun to differentiate motivation domains in schizophrenia, with many (c.f., J. Choi, Mogami, & Medalia, 2010; Gard et al., 2014; Medalia & Brekke, 2010) looking to Self-Determination Theory (SDT; Deci & Ryan, 1985a, Ryan and Deci, 2000a) to conceptualize the specific domains of motivation in schizophrenia. SDT posits that there are three types of motivation differentiated by the underlying reason for a behavior: intrinsic motivation, extrinsic motivation, and amotivation. Given the prominence of SDT and these motivation domains in schizophrenia research, each of these domains will be briefly reviewed.

Intrinsic motivation, which is often defined as pursuing a task or action because it is enjoyable or interesting (Ryan & Deci, 2000a), has recently received considerable attention in schizophrenia research in part because of its associations with cognitive performance (Fervaha, Zakzanis, et al., 2014) and functioning (Luther, Firmin, Vohs, et al., 2016). Importantly, clinician-rated intrinsic motivation has been found to be overlapping but distinct from clinician-rated amotivation and overall negative symptoms (Luther et al., 2015; Yamada et al., 2010). However, somewhat surprisingly, self-reported intrinsic motivation measures frequently demonstrate small correlations with clinician-rated intrinsic motivation or negative symptom measures (Barch, Yodkovik, Sypher-Locke, & Hanewinkel, 2008; Breitborde, Woolverton,

Frost, & Kiewel, 2014; J. Choi et al., 2014), leading to calls for additional evaluation of the overlap between extant motivation measures (J. Choi et al., 2014; Cooper et al., 2015).

Extrinsic motivation, or pursuing a task or action in order to obtain an outcome different from the task itself, such as an external reward (i.e., monetary reward or praise) or to avoid punishment (Ryan & Deci, 2000a), has long been an area of interest in psychotherapeutic treatment in schizophrenia. Indeed, the use of rewards has been a main component used to facilitate skill attainment, treatment engagement, learning, and improved behavior in token economies (Werry, 1969) and social learning programs (Glynn & Mueser, 1986) for people with schizophrenia. More recently, a burgeoning area of research has found that people with schizophrenia demonstrate impaired external reward processing in the areas of value representations—identifying and updating mental representations of reward value associated with a task (Gold, Waltz, Prentice, Morris, & Heerey, 2008; Strauss, Waltz, & Gold, 2014)—and effort-cost computations or estimating whether the benefits outweigh the effort or costs associated with completing a task (Gold et al., 2013; Green, Horan, Barch, & Gold, 2015). These findings have led researchers to posit that these reward processing deficits are key processes underlying motivation deficits in schizophrenia (Gold et al., 2008; Strauss et al., 2014). However, studies examining the association between performance on reward processing tasks (e.g., effort-based decision-making tasks) and clinician-rated amotivation or negative symptoms have revealed mixed findings (c.f., McCarthy et al., 2016; Treadway et al., 2015). One reason for these incongruent findings is that many studies consisted of small samples (e.g., 16 people; Fervaha, Graff-Guerrero, et al., 2013), limiting their power to detect significant associations. By examining associations across multiple studies, a comprehensive meta-analytic approach can

offset small sample sizes and may offer additional clarity regarding the association between these measures.

Finally, amotivation, or lacking the impetus or intention to carry out a task or action (Ryan & Deci, 2000a), has long been an area of interest in schizophrenia (Bleuler, 1911/1950; Kraepelin, 1913/1919). Amotivation has been conceptualized as one of two key subdomains of negative symptoms (American Psychiatric Association, 2013; Marder & Galderisi, 2017) and is related to, but distinct from, the second negative symptom domain of diminished expression symptoms (i.e., blunted affect, alogia). Moreover, these negative symptom domains have demonstrated differential associations with other clinical variables, with some research showing that amotivation is more strongly linked to both functioning and neurocognition than diminished expression negative symptoms (Green, Hellemann, Horan, Lee, & Wynn, 2012; Strauss et al., 2013).

### **Measuring Motivation in Schizophrenia**

**Self-reported motivation measures.** Numerous self-reported measures have been used to assess these three motivation domains. One of the most recent and widely used self-reported motivation measures is the Intrinsic Motivation Inventory for Schizophrenia Research (IMI-SR; J. Choi et al., 2010), which is derived from SDT and includes subscales of interest/enjoyment, effort, and competence. The IMI-SR was developed to measure state intrinsic motivation for a particular task—most frequently the IMI-SR is used to measure intrinsic motivation for cognitive remediation or computerized learning tasks but has also been used by Cooper et al. (2015) to measure participants' intrinsic motivation for “an important task they recently completed” (Cooper et al., 2015, p. 72). The IMI-SR has gained notable traction, in part because it was one of the first non-negative symptom (i.e., amotivation) measures of motivation to be validated in a

schizophrenia sample. Further, although some have expressed concerns about obtaining reliable data on self-reported measures from people with schizophrenia (Gupta, Holhausen, Gou, & Bowie, 2014; Takeuchi, Fervaha, & Remington, 2016), others have posited that because intrinsic motivation reflects internal states (i.e., the enjoyment one derives from an activity or the reason behind completing an activity), intrinsic motivation can only be truly measured with a self-reported measure (J. Choi et al., 2014; Kremen et al., 2016).

More recently, Cooper and colleagues (2015) modified and validated a clinical version of the General Causality Orientation Scale (GCOS; Deci & Ryan, 1985b) for use with people with schizophrenia and other severe mental illnesses (GCOS-clinical populations: GCOS-CP). The GCOS is another SDT-derived measure that uses vignettes to assess three trait-like motivation orientations—relatively enduring aspects of personality that underlie how a person generally initiates and regulates behavior (Deci & Ryan, 1985b). After reading each vignette, participants are asked to imagine themselves in that situation and rate several items based on how they would likely respond. The three GCOS motivation orientations assessed are autonomy, control, and impersonal, which map onto the domains of intrinsic motivation, extrinsic motivation, and amotivation, respectively. However, studies examining the association between the GCOS, GCOS-CP, and IMI-SR and clinician-rated negative symptoms have demonstrated mixed findings, with some finding small correlations and others finding medium correlations (c.f., J. Choi et al., 2014; J. Choi, Fiszdon, & Medalia, 2010; Cooper et al., 2015). Importantly, however, most prior studies examined the association between these scales and a broader range of negative symptoms (i.e., total scores); thus, examining the association between these scales and the amotivation negative symptom domain can provide a more precise test of the convergent validity of these scales.

Other self-reported measures of motivation include negative symptom self-reported scales that assess amotivation. Lincoln, Dollfus, and Lyne (2017) recently conducted a comprehensive review of negative symptom measures and identified six negative symptom self-reported measures that assess amotivation. First, most recently, researchers have used the Motivation and Pleasure Scale-Self Report (MAP-SR; Llerena et al., 2013), which is the self-report counterpart of the motivation and pleasure subscales of the Clinical Assessment Interview for Negative Symptoms (CAINS; Horan, Kring, Gur, Reise, & Blanchard, 2011; Kring, Gur, Blanchard, Horan, & Reise, 2013). The MAP-SR contains a subscale that assesses participants' motivation and effort to engage in activities with other people, school or work activities, and hobbies or recreational activities over the past week. Notably, scores on the MAP-SR have shown large correlations with clinician-rated measures of negative symptoms (Engel & Lincoln, 2016; Llerena et al., 2013), but the convergent validity with the individual amotivation subscales on negative symptom measures has not been assessed. Other negative symptom self-reported measures that assess amotivation include the Subjective Experience of Deficits in Schizophrenia (SEDS; Liddle & Barnes, 1988), which contains a subscale that assesses a person's motivation and energy, and the Subjective Experience of Negative Symptoms (SENS; Selten, Sijben, Van den Bosch, Omloo-Visser, & Warmerdam, 1993) scale that assesses respondents' subjective experience of the symptoms, including the avolition-apathy subscale, on the Scale for the Assessment of Negative Symptoms (SANS; Andreasen, 1984). However, the convergent validity of the motivation subscales of several of the older self-reported measures such as the SEDS and SENS has not been adequately explored with other measures of motivation (Lincoln et al., 2017).

**Clinician-rated motivation measures.** Measures of motivation that are clinician-rated largely assess amotivation, apathy, or avolition—terms often used interchangeably in the



schizophrenia literature to describe negative symptom subscales related to reduced motivation (c.f., Foussias & Remington, 2010; Lincoln et al., 2017; Markou et al., 2013). As Lincoln et al. (2017) noted, almost all of the most widely used negative symptom measures contain subscales that assess reductions in motivation. For example, the SANS contains the avolition-apathy subscale that assesses participants' energy, drive, and interest for work and/or school, daily self-care, and non-passive activities, and the Brief Negative Symptom Scale (BNSS; Kirkpatrick et al., 2011; Strauss et al., 2012) contains an avolition subscale that assesses participants' initiation and persistence for hobbies, self-care, treatment, and general activities.

Aside from amotivation subscales from negative symptom measures, many schizophrenia researchers have also used two other clinician-rated measures to measure motivation. The first is the widely used three-item intrinsic motivation index from the clinician-rated Heinrichs-Carpenter Quality of Life Scale (Heinrichs, Hanlon, & Carpenter, 1984), which was originally put forth by Nakagami, Xie, Hoe, and Brekke (2008). The index assesses participants' sense of purpose, curiosity, and degree of motivation. The second measure, the clinician-rated version of the Apathy Evaluation Scale (AES-C; Marin, Biedrzycki, & Firinciogullari, 1991), has recently received increased attention in part due to the fact that it was developed specifically to measure apathy/amotivation. The original clinician-version of the AES (AES-C) contains 18-items and was introduced as a single factor measure that could be used for a range of clinical populations to assess motivational loss that was due to concurrent decreases in "the overt behavioral, cognitive, and emotional concomitants of goal-directed behavior" (Marin et al., 1991, p. 143). Faerden et al. (2008) conducted a factor analysis of the AES-C in a first-episode psychosis sample, finding that an abbreviated 12-item version of the AES-C provides a slightly more valid and reliable assessment of apathy in first-episode psychosis than the 18-item measure. Further, scores on both

the AES-C (12 and 18-item versions) and the QLS intrinsic motivation index have demonstrated adequate reliability as well as convergent validity with clinician-rated motivation or overall negative symptom scales (J. Choi et al., 2014; Faerden et al., 2008, 2009; Luther et al., 2015), but the convergent validity between these measures and self-reported and performance-based measures of motivation is not as established.

**Performance-based measures.** Interest in creating more objective measures of motivation has led to the development and adaption of several performance-based paradigms to measure motivation (Fervaha, Duncan, et al., 2015; Green, et al., 2015), specifically extrinsic motivation, in schizophrenia. One difficulty in devising such assessments is that motivation is an important factor for completing most performance-based tasks (Markou et al., 2013). Further, performance-based tasks used to measure motivation often involve other processes such as reinforcement learning (i.e., modifying behavior in response to positive or negative feedback). To address these concerns, the Motivation Working Group of the Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia (CNTRICS; Carter et al., 2008) initiative identified and differentiated translational performance-based tasks that assess the most fundamental and unconfounded processes related to either motivation or reinforcement learning (Markou et al., 2013). This resulted in the identification of two types of motivation tasks, which both capture extrinsic motivation: Effort-based decision-making and outcome devaluation tasks. Notably, the selection of tasks was guided by the availability of evidence from human and animal research supporting the construct validity of each task. Accordingly, in order to reduce construct validity concerns (i.e., including measures that involve disparate processes unrelated to motivation), only measures that align with these two types of performance-based motivation measures were eligible for inclusion in the current meta-analyses. However, although some

studies have used outcome devaluation tasks with psychiatric populations (Gillan et al., 2011), these tasks are still primarily used in non-human studies and to our knowledge have not yet been used in people with schizophrenia; thus, only effort-based decision-making tasks were included in the current study.

Effort-based decision-making tasks generally assess how much effort a person is willing to exert for a given level of reward. On these tasks, which have been described by Green et al., (2015) and Reddy, Horan & Green (2015), participants generally complete a computerized task that has forced choice-trials where participants are asked to choose between an option that requires low effort and provides low rewards or an option that requires relatively more effort and provides greater rewards. Several tasks assessing effort-based decision-making have been adapted from paradigms used to measure motivation with animals. Among these, the most widely used is the Effort Expenditure for Rewards Task (EEfRT; Treadway, Buckholtz, Schwartzman, Lambert, & Zald, 2009), which is a computerized effort-based decision-making task where participants choose between performing an easy task (pressing a computer key 30 times in seven seconds with their index finger on their dominant hand) or a hard task (pressing a computer key 100 times in 21 seconds using their little finger of their non-dominant hand) for a specified amount of monetary reward on each trial; easy trials are worth \$1.00, while the hard trials range in value from \$1.24 to \$4.30. Additionally, the probability of receiving the monetary award is also specified before each trial, with the probability of winning the specified award ranging from low (12%), medium (50%), and high (88%). After each trial, participants receive feedback about whether they won the reward.

Other effort-based tasks used with people with schizophrenia have also been described by Green et al. (2015) and Reddy, Horan, & Green (2015) and generally involve forced choice-trials

where effort can be physical (i.e., squeezing a handgrip) or cognitive (i.e., identifying numbers on cards as odd or even) and increased effort is associated with greater rewards that are generally monetary. However, as others have noted (Green et al., 2015; Horan et al., 2015; Reddy, Horan, & Green 2015), many effort-based decision-making tasks have not received adequate psychometric evaluation in people with schizophrenia, and studies examining the relationship between effort-based decision-making tasks and clinician-rated motivation or negative symptom measures have yielded mixed results (c.f., McCarthy, Treadway, Bennett, & Blanchard, 2016; Treadway, Peterman, Zald, & Park, 2015). A meta-analytic approach may help to clarify these associations. Further, there is no agreed upon score derived from these measures that is used to assess motivation (Horan et al., 2015), which can complicate comparisons between studies (Reddy, Horan, & Green, 2015); therefore, to reduce the heterogeneity among scores for the current investigation, when appropriate, the most commonly used motivation score of percent of hard choices chosen across trials was used for the effort-based-decision-making tasks.

Notably, effort-based decision-making tasks also differ from reinforcement learning tasks that are used in schizophrenia research. Reinforcement learning tasks generally ask a participant to repeatedly complete a task (i.e., pushing a button at a certain time or choosing one of two geometric shapes), which is followed by monetary (You won/lost \$0.20) and/or verbal feedback (“Correct” or “Incorrect”) that is delivered probabilistically and varies in magnitude across the trials (Insel et al., 2014; Strauss et al., 2011). For these tasks, participants are instructed to incorporate feedback and adjust their behavior on the following trial in order to maximize verbal or monetary reward receipt. Thus, although motivation is involved in the completion of reinforcement learning tasks, these tasks differ in that they require participants to learn the appropriate response and then continuously update their responses based on feedback, making it

difficult to precisely identify if reduced task performance is due to difficulties in reinforcement learning or motivation. Thus, reinforcement learning tasks were not included in the current investigation.

### **Current Study**

The overall aim of the current study was to conduct a series of meta-analyses that summarize the relationships between self-reported, clinician-rated, and performance-based measures of motivation across existing studies of people with schizophrenia-spectrum disorders. Given that SDT (Deci & Ryan, 1985a) has been a prominent organizing framework for motivation in schizophrenia research to date, the conceptualization proposed by Ryan and Deci (2000a) was utilized to identify and categorize motivation measures into intrinsic motivation, extrinsic motivation, or amotivation. Inclusion of measures that fall within these domains not only aligns with current thinking of motivation in schizophrenia but also will reduce the challenges in operationalizing motivation and help to alleviate construct validity concerns. Specifically, for self-reported and clinician-rated measures of intrinsic motivation and extrinsic motivation, we used a recent qualitative review by Kremen, Fiszdon, Kurtz, Silverstein, & Choi (2016) that describes extant intrinsic motivation and extrinsic motivation measures to guide measure inclusion. For self-reported and clinician-rated measures of amotivation, we included measures described as measuring amotivation in the review by Lincoln et al. (2017). These self-reported and clinician-rated measures were chosen because they are among the most widely used and have been validated in people with schizophrenia, which can also help to reduce construct validity concerns. Lastly, the performance-based measures that were included, which all assess extrinsic motivation, were measures that aligned with Markou et al.'s (2013) review.

The specific aims were to:

- 1) Identify the magnitude of the association between self-reported, clinician-rated, and performance-based measures of motivation. Consistent with some prior research (c.f., Evensen et al., 2009; Faerden et al., 2009; Strauss et al., 2016) and the notion that these measures are putatively assessing the same construct, we hypothesized that the magnitude of the relationship between these three motivation measurement types would be significant, positive, and represent a medium effect size.
- 2) Identify the magnitude of the associations across measurement types *within* motivation domains (e.g., association between self-reported and clinician-rated intrinsic motivation). We hypothesized that there would be significant, medium, and positive relationships within the same motivation domain across the different assessment types.
- 3) Examine moderators of the relationships between the different motivation measurement types, including age, gender, percent of the sample with a schizophrenia diagnosis, length of illness, and chlorpromazine equivalent doses. Given that shorter length of illness (Luther et al., 2015), and lower chlorpromazine equivalent doses (Kirsch, Ronshausen, Mier, & Gallhofer, 2007; Luther, Firmin, Minor, et al., 2016) have been linked to higher motivation, we hypothesized that the relationships between the motivation measurement types will be stronger with decreasing length of illness and chlorpromazine equivalent doses. All other moderator analyses were exploratory.

## Method

### Meta-Analytic Reporting Guidelines

The evidence-based Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher, Liberati, Tetzlaff, & Altman, 2009) guidelines were used to ensure quality reporting of the meta-analytic findings.

## Literature Search

After consulting with a scientific librarian and referencing previously published meta-analyses (O'Driscoll, Liang, & Mason, 2014; Pearlman & Najjar, 2014; Salyers et al., 2016) and guidelines (Card, 2012; Lipsey & Wilson, 2001), we conducted three steps to identify studies. First, searches were conducted in *Embase*, *Medline*, *PsychINFO*, *PsychARTICLES*, *Pubmed*, and *Web of Science Core Collection*, covering articles, theses/dissertations, and conference proceedings that were published up to September 7th, 2017. We included conference proceedings and theses/dissertations in order to reduce the possibility of publication bias and the file drawer problem (i.e., non-significant relationships not getting published in peer-reviewed articles) (Borenstein et al., 2009; Lipsey & Wilson 2001). Including conference abstracts was also based on the recommendations of several meta-analytic experts (Card, 2012; Lipsey & Wilson, 2001) and guidelines and standards for meta-analytic and systematic reviews from Cochrane (Higgins & Greene, 2011), the Agency for HealthCare Research and Quality (Balshem et al., 2013), and the Institute of Medicine (Eden, Levit, Berg, & Morton, 2011). Keywords for the search included all possible derivatives of the terms motivation and schizophrenia as well as an extensive combination of terms pertaining to measurement types (e.g., self-reported, clinician-rated, performance-based). When available, English language and human filters were used. Second, reference sections of seven qualitative reviews related to the conceptualization or measurement of motivation in schizophrenia were reviewed (i.e., Green & Horan, 2015; Green et al., 2015; Kremen et al., 2016; Markou et al., 2013; Medalia & Brekke, 2010; Reddy, Horan, &

Green 2015; Strauss et al., 2014). Third, given that articles needed to contain two different types of motivation measures, we conducted forward searches of the original and/or psychometric validation articles for the included self-reported and performance-based motivation measures. See Figure 1 for PRISMA flow chart diagram.

### **Study Selection: Inclusion and exclusion criteria**

Studies were eligible if they were 1) available in English, 2) included an empirical measure of at least two of the following: a self-reported, clinician-rated, or performance-based measure of motivation and, 3) included at least two motivation measures that mapped onto the domains of intrinsic motivation, extrinsic motivation, or amotivation. As described above, this process was grounded in extant literature and thus was guided by the measures described in the reviews by Kremen et al. (2016), Lincoln et al. (2017), and Markou et al. (2013). Specifically, we only included intrinsic motivation measures that aligned with those described by Kremen et al. (2016). Eligible amotivation measures were those that were described by Lincoln et al. (2017). For extrinsic motivation, all the performance-based motivation measures that aligned with those described by Markou et al. (2013) were eligible, as they assess a person's willingness to exert effort for an external reward. In addition, self-reported measures of extrinsic motivation were also eligible to be included if they aligned with those described by Kremen et al. (2016). All potentially eligible measures that were considered for inclusion are reported in Appendix Table 1. Additional eligibility criteria included that studies 4) assessed motivation in a schizophrenia-spectrum sample and 5) reported a bivariate relationship between the included motivation measures. Only baseline relationships were examined in intervention or experimental studies. Studies were excluded if bivariate relationships were not available after contacting the authors or if a study sample overlapped with the sample from another study.



## Coding

Eligible studies were coded according to a codebook that was developed based on guidelines from Card (2012) and Lipsey and Wilson (2001).

**Sample-level information.** Sample size and mean age, length of illness, chlorpromazine equivalent doses, as well as gender and diagnosis (e.g., percent schizophrenia), were coded for each study. Measurement type (i.e., self-reported) and motivation domain assessed were also coded.

**Effect size-level coding.** For each study, the raw effect size (i.e., correlation coefficient) was extracted, representing the relationship between motivation measurement types. All effect sizes were converted into Pearson's correlations. Further, for the measurement meta-analyses, if necessary, the effect size was reverse coded so that higher values on each motivation measure represented greater motivation. For the domain meta-analyses, if necessary, the effect sizes were reverse coded so that all domain measures were scored in the same direction (i.e., higher scores = greater amotivation, greater intrinsic motivation, and greater extrinsic motivation). If a study included multiple effect sizes that captured the same type of relationship (i.e., relationship between self-reported and clinician-rated motivation measures) or domains (i.e., self-reported and clinician-rated intrinsic motivation), the effect sizes assessing the same type of relationship or domain were averaged and weighted by sample size in order to reduce bias and not violate the assumption of independence (Card, 2012). For both the measurement and the domain meta-analyses, if a measure assessed multiple symptom or motivation domains, we included only relationships with the subscale or items that aligned with motivation or the respective domain of interest when coding and ultimately analyzing the effect sizes. If a study did not report the relevant correlation or measure information, study authors were emailed. All data were originally

coded into Excel and then imported into the Statistical Package for the Social Sciences (SPSS), version 22.0 and Comprehensive Meta-Analysis, Version 3 (CMA; Borenstein et al., 2014).

### **Analyses**

Descriptive statistics were conducted using SPSS. Mean overall effect sizes were then computed in CMA using a random-effects model, which accounts for both within-study and between-study variability (Lipsey & Wilson, 2001). We first aimed to conduct three separate meta-analyses to investigate the relationships between 1) self-reported and clinician-rated, 2) performance-based and clinician-rated, and 3) self-reported and performance-based motivation measures. Based on extant motivation measures described in Kremen et al. (2016), Lincoln et al. (2017), and Markou et al. (2013), we also aimed to calculate three additional meta-analyses examining the relationships within motivation domains across the different assessment types: 1) self-reported and clinician-rated intrinsic motivation, 2) self-reported and clinician-rated amotivation, and 3) self-reported and performance-based extrinsic motivation measures. For all meta-analyses, the magnitude of the overall effect sizes was interpreted based on Cohen's (1992) recommendation for correlations where .10 is small, .30 is medium, and .50 is large.

For each meta-analysis, a one-study removed sensitivity analysis was conducted in CMA to determine if one study was unduly impacting the overall effect size. This analysis is conducted by running each meta-analysis repeatedly, with a different study removed each time (Borenstein, Hedges, Higgins, & Rothstein, 2009). CMA produces a forest plot and effect size point estimates that visually and numerically depict how the overall mean effect size would be affected if each study were removed one at a time (Borenstein et al., 2009). Studies were considered for removal if they visually appeared to be outliers and the overall effect size substantially changed after removing the study.

The presence of heterogeneity among the included effect sizes of each meta-analysis was identified using the  $Q$ -statistic (Card, 2012). To assess the extent of any identified heterogeneity, the  $I^2$  index was examined (Higgins & Thompson, 2002; Huedo-Medina, Sánchez-Meca, Marín-Martínez, & Botella, 2006) to determine the percentage of variation that is due to between-study variability; values greater than or equal to 25% suggest greater between-study variability than would be expected by chance (Huedo-Medina et al., 2006). More specifically,  $I^2$  index values of 25%, 50% and 75% are considered low, medium, and high amounts of variability, respectively (Higgins & Thompson, 2002; Huedo-Medina et al., 2006). Moderation analyses for each meta-analysis were conducted when the  $Q$ -statistic was significant and the  $I^2$  index was 25% or greater.

All proposed moderators were continuous and were assessed in CMA with meta-regressions using a random effects model. At least six studies needed to provide data in order for each moderation analysis to be conducted (Fu et al., 2011). Because meta-regressions use listwise deletion, each moderator was examined individually to maximize the number of studies included. Moderators were considered significant if the associated beta weight was significant ( $p < .05$ ) and the  $I^2$  index decreased when compared to the  $I^2$  index of the main corresponding meta-analysis.

Finally, publication bias was also examined using two steps. First, funnel plots were used to see if they were roughly triangular in shape or had an asymmetrical distribution around the mean effect size (Borenstein et al., 2009; Card, 2012), which would indicate potential bias. Second, Egger's regression test (Egger, Smith, Schneider, & Minder, 1997), which regresses the normalized effect estimates (i.e., effect size divided by its standard error) against studies' precision (Egger et al., 1997), was conducted for each meta-analysis; Egger's test suggests that

publication bias is present when the intercept is significant ( $p < .05$ ). Some authors suggest that a minimum of 10 studies is needed in order to be adequately powered to detect publication bias (Kepes, Banks, McDaniel, & Whetzel, 2012).

## Results

### Study Characteristics

Forty-five unique studies were eligible to be included in the meta-analyses (see Figure 1 for Study Retrieval Flow Diagram). Of these, 33 examined the relationship between self-reported and clinician-rated motivation measures, 12 examined the relationship between performance-based and clinician-rated motivation, and two studies assessed the relationship between self-reported and performance-based motivation measures (one study assessed all three relationships). Summary study characteristics appear in Table 1, and individual study characteristics, motivation measures, and effect sizes are presented in Appendix Table 2. Across all eligible studies, 2,781 participants with a schizophrenia-spectrum diagnosis were included, representing 13 countries. The most commonly used self-reported measure was a version of the Intrinsic Motivation Inventory for Schizophrenia Research (IMI-SR; J. Choi et al., 2010), while the most commonly used clinician-rated motivation measure was the amotivation subscale (Fervaha, Foussias, Agid, & Remington, 2014; Liemburg et al., 2013) of the Positive and Negative Syndrome Scale (PANSS; Kay et al., 1987). Finally, the majority of studies measuring performance-based motivation used the Effort Expenditure for Rewards Task (Treadway, Buckholtz, Schwartzman, Lambert, & Zald, 2009). Forty-two studies had been published in a peer-reviewed journal article, while three studies were conference abstracts.

For the domain meta-analyses, four studies assessed the relationship between self-reported and clinician-rated intrinsic motivation, and 23 assessed the relationship between self-

reported and clinician-rated amotivation. No studies assessed the relationship between self-reported and performance-based extrinsic motivation. See Table 2 and Appendix Table 2 for more information.

Of note, several of the proposed meta-analyses had few studies (i.e., 2 or 4). However, based on Borenstein et al.'s (2009) recommendation to conduct and report meta-analyses even when the number of studies is small and in light of our goal to comprehensively review the overlap between different motivation measurement types in extant studies, we still conducted meta-analyses with few studies even though they were likely underpowered to detect effects.

### **Sensitivity Analyses**

Visually examining the one-study removed forest plots (available from the first-author upon request) for each of the five meta-analyses (three measurement type and two domain) suggested that there was some variation among the study-level effect sizes. In particular, the meta-analysis examining the relationship between performance-based and clinician-rated motivation appeared to contain a potential outlier. Examination of the effect size point estimates suggested that McCarthy, Treadway, Bennett, & Blanchard (2016) may be overly influencing the mean effect size; compared to most of the other included studies, this study also had several demographic differences (e.g., older sample). Given the effect size impact and demographic differences of McCarthy et al. (2016), it was removed from all further analyses. No additional studies were removed from the remaining meta-analyses.

### **Main Analyses**

For the meta-analyses examining the overall associations between different motivation measurement types, two of the three meta-analyses produced significant effect sizes, which were positive and small to medium in magnitude; only the meta-analysis between self-reported and

performance-based motivation measures was non-significant, evidencing a negligible overall effect size. Of the two conducted meta-analyses examining the relationships within motivation domains across different measurement types, only the meta-analysis examining the association between self-reported and clinician-rated amotivation was significant, demonstrating a medium effect size. A summary of the meta-analytic statistics is presented in Table 2.

**Self-reported—clinician-rated meta-analysis.** Thirty-three studies (including the three conference abstracts) had effect size data for the meta-analysis examining the relationship between self-reported and clinician-rated motivation measures (see Table 2, Figure 2). Generally consistent with our hypothesis, there was a positive, significant overall effect size between these motivation measures that was approaching a medium effect size ( $r = .27, p < .001$ ), suggesting that higher levels of self-reported motivation were associated with higher levels of clinician-rated motivation. Notably, this correlation suggests that the shared variance between these motivation measurement types was approximately 7.29%. Further, sample-level correlations ranged from  $r = -.11$  to  $r = .75$ . To ensure that DeRosse, Nitzburg, Kompancaril, and Malhotra (2014) was not unduly influencing the effect size given that this study made up 21.4% of the overall sample, we also calculated the effect size without this study; this revealed a similar effect size ( $r = .28, p < .001$ ). Similarly, we also assessed the impact of the conference abstracts on the overall effect size; when the three conference abstracts were excluded, the effect size was again similar ( $r = .28, p < .001$ ), suggesting that the conference abstracts were also not unduly influencing the effect size. Thus, both DeRosse et al. (2014) and the conference abstracts were retained in all additional analyses. In terms of heterogeneity, the  $Q$ -statistic for the overall effect size with all 33 studies was significant ( $p < .001$ ), with the  $I^2$  index (72.70%) suggesting there was a medium to high amount of heterogeneity and that moderator analyses would be appropriate (see below).

**Performance-based—clinician-rated meta-analysis.** After removing McCarthy et al. (2016), there were 11 studies that provided data for the relationship between performance-based and clinician-rated motivation measures (Table 2, Figure 3). Analyses indicated that there was a positive, small, significant overall effect size ( $r = .21$ ;  $p < .001$ ) between these two measurement types, suggesting that higher levels of performance-based motivation were associated with higher levels of clinician-rated motivation. Of note, this finding is in the hypothesized direction, but the magnitude of the effect size was slightly lower than hypothesized. Further, this correlation suggests that the shared variance between these motivation measurement types was approximately 4.41%. All of the included samples found positive relationships between the measurement types (ranging from  $r = .05$  to  $r = .52$ ). Significant heterogeneity was not present among the effect sizes ( $Q$ -statistic,  $p = .32$ ;  $I^2 = 13.09\%$ ).

**Self-reported—Performance-based motivation association.** Two studies provided data on the relationship between self-reported and performance-based motivation (Table 2, Figure 4). Contrary to hypothesis, the overall effect size was negligible and non-significant ( $r = -.001$ ,  $p = .992$ ). One of the studies found a negative relationship ( $r = -.07$ ), while the other study found a positive relationship between these measurement types ( $r = .17$ ). There was not significant heterogeneity in the included effect sizes ( $Q$ -statistic,  $p = .26$ ;  $I^2 = 20.71\%$ ). However, given that these analyses were based on only two studies, the results should be interpreted cautiously.

**Motivation domain associations.** Two of the within domain meta-analyses were able to be conducted (see Table 2). Consistent with our hypothesis, the overall effect size for the meta-analysis between self-reported and clinician-rated amotivation was significant, positive, and medium in magnitude ( $r = .34$ ,  $p < .001$ ); this was based on 23 studies. The effect size for the meta-analysis examining the relationship between self-reported and clinician-rated intrinsic

motivation was small and positive but non-significant ( $r = .16, p = .259$ ); however, because only four studies were included, this meta-analysis was likely underpowered. See Figures 5 and 6 for forest plots. Heterogeneity analyses indicated that significant heterogeneity was present to a medium to large extent among the associations between both self-reported and clinician-rated amotivation ( $Q$ -statistic,  $p < .001$ ;  $I^2 = 77.34\%$ ) and intrinsic motivation measures ( $Q$ -statistic,  $p = .008$ ;  $I^2 = 74.67\%$ ).

### **Moderator Analyses**

**Meta-regression analyses.** Given that significant heterogeneity was present to a medium or large extent for three of the meta-analyses, we aimed to conduct moderator analyses in attempts to identify the source of the heterogeneity in these meta-analyses. However, the meta-analysis of self-reported and clinician-rated intrinsic motivation had only four studies (i.e., did not meet the threshold of six; Fu et al., 2011), making it ineligible for moderator analyses. We examined five potential continuous moderators, including mean age, length of illness, chlorpromazine equivalent doses, as well as % female and % schizophrenia in the meta-analyses between self-reported and clinician-rated motivation measures and then self-reported and clinician-rated amotivation measures. Results are presented in Table 3. Consistent with our hypothesis, length of illness significantly moderated the relationship between self-reported and clinician-rated motivation ( $b = -.014, p = .04, I^2 = 59.26$ ) so that for every one year increase in length of illness, the relationship between self-reported and clinician-rated motivation weakened by .014. Contrary to our hypothesis, mean chlorpromazine equivalent doses were not significant moderators of either relationship. No additional moderators were significant, and there were no significant moderators of the relationship between self-reported and clinician-rated amotivation.

### **Publication Bias**



Analyses examining publication bias were conducted for all meta-analyses with more than two studies (i.e., a minimum of three studies is required to conduct the analyses) to ensure that bias was not influencing both significant and non-significant overall mean effect sizes. Funnel plots (available from the first-author upon request) of the effect sizes for the four eligible meta-analyses were relatively symmetrical and triangular, suggesting that publication bias was not present. Egger's regression test of the intercept was also not significant (all  $p$ 's > .05) for any of the four meta-analyses, further supporting the notion that publication bias was not present. However, given that several of the meta-analyses had fewer than or close to 10 studies, it is possible that we were underpowered to detect publication bias in these meta-analyses.

### Discussion

In an effort to inform the conceptualization, measurement, and treatment of motivation in schizophrenia, we conducted several meta-analyses to identify the overlap or discrepancy between extant motivation measures. Forty-five unique studies were identified, and mean weighted overall effect sizes were calculated between clinician-rated and both self-reported and performance-based motivation measures and between self-reported and performance-based motivation measures. Partially consistent with our hypotheses, our results suggest that there are significant and positive relationships between clinician-rated and both self-reported and performance-based motivation measures that represent small to medium overall effect sizes. Further, these results appeared not to be unduly influenced by outliers or publication bias. We also found a negligible, non-significant overall effect size between self-reported and performance-based motivation measures, which was contrary to our hypothesis. However, given that only two studies were included in the self-reported and performance-based motivation

measures meta-analysis, these results should be interpreted cautiously, as we were likely underpowered to detect effects.

The overall effect size between self-reported and clinician-rated motivation measures was also relatively larger than the overall effect sizes between performance-based motivation and both clinician-rated and self-reported motivation measures. This may in part be explained by the greater degree of similarity between self-reported and clinician-rated assessment methods. First, the self-reported and clinician-rated motivation measures often have similar item content. For example, both the MAP-SR and SANS contain items assessing motivation for work and/or school. Second, both self-reported and clinician-rated motivation measures also have a shared data source: the participants' self-reported information. Although clinician-rated measures may be more objective measures of motivation as they incorporate divergent sources of information (e.g., medical chart or informant information) and clinician judgment (e.g., determining the validity of the participant's self-reported information), the shared data source could also help to explain the stronger observed correlation.

In contrast, the performance-based measures do not share method variance and have less content overlap with both the clinician-rated and self-reported motivation measures; their relationships were likely smaller due to several important factors. First, performance-based measures assess one's behavioral performance on a time-limited task (in a lab setting). On the other hand, clinician-rated measures utilize a combination of self-reported, clinician assessed, and informant reported perceptions of behavior for a longer time period (i.e., week to a month) out in the community, while self-reported measures focus solely on a person's perceptions of their behavior generally over the past week or month. Further, although all broadly assess one's willingness to expend effort or to complete a task (i.e., pushing a button vs. going to work or

school), the time frame of reward receipt and the type of tasks/level of effort involved differ. For example, within the performance-based measures, participants are reinforced almost instantaneously and frequently during the task, whereas clinician-rated and self-reported measures largely assess domains that involve delayed rewards (e.g., a paycheck or an educational degree). This is particularly important given that people with schizophrenia have difficulties representing, maintaining, and updating the value of future rewards (Strauss et al., 2014), especially as the length of time of reward receipt increases (Heerey, Robinson, McMahon, & Gold, 2007); thus, we might expect divergent behavioral responses (i.e., greater or reduced effort or task engagement) to tasks that offer frequent rewards versus tasks that offer less frequent rewards. Similarly, considerably more effort is required to work or attend school compared to pushing a button or gripping a lever. Thus, the divergent timeframe, rewards, and tasks assessed likely contribute to the smaller observed effect sizes between performance-based and both clinician-rated and self-reported motivation measures.

Of note, our findings also point to the need for additional work aimed at clarifying the overlap or disparity between self-reported and performance-based motivation measures. Only two studies were eligible to be included in the meta-analysis examining this relationship, and the effects in these studies were in opposite directions. Considering that prior studies have questioned the validity and reliability of self-reported measures for people with schizophrenia (Gupta, Holshausen, Gou, & Bowie, 2014; Takeuchi, Fervaha, & Remington, 2016) and several researchers have pointed to the need for additional psychometric evaluation studies of performance-based motivation measures (Barch, Gold, & Kring, 2017; Reddy et al., 2015), future studies are needed to more conclusively validate these measures in schizophrenia samples

and understand the level of overlap in the underlying construct(s) that both self-reported and performance-based motivation measures are assessing.

Despite these differences in methods of measurement, the magnitude of the effect sizes still suggests some evidence of convergent validity, particularly between clinician-rated and both self-reported and performance-based motivation measurement methods. However, compared to the effect size between self-reported and clinician-rated motivation measures, there was relatively less convergent validity between the performance-based and clinician-rated motivation measures. Importantly, although there is some evidence of convergent validity between clinician-rated and both self-reported and performance-based motivation measures, the magnitude of the overall effect sizes also suggests that these measurement types are assessing small amounts of shared variance. Indeed, the overall effect size between self-reported and clinician-rated motivation measures indicated that the measures were assessing only 7.29% of similar underlying variance, while the performance-based and clinician-rated measures were sharing only 4.41% of similar variance. These findings indicate that although clinician-rated and both self-reported and performance based-measures are assessing some shared underlying construct, there is more variability than similarity in the underlying construct that is being assessed by the different motivation measurement types.

When possible, we also conducted moderator analyses among the meta-analyses with significant heterogeneity, namely the overall self-reported and clinician-rated meta-analysis as well as the self-reported and clinician-rated intrinsic motivation and amotivation domain meta-analyses. The remaining two meta-analyses did not have significant heterogeneity, suggesting there was more similarity among the included study effect sizes. Consistent with our hypothesis, length of illness was a significant moderator of the overall relationship between self-reported and

clinician-rated motivation measures, so that as length of illness increased, the relationship between the motivation measures weakened. Although the exact mechanism causing this reduced relationship as the length of illness progresses is not evident from the current findings, there are several factors that could be explored in future research. First, as the length of illness increases and motivation deficits become more prolonged, some areas or domains of motivation might be differentially impacted over time (Luther et al., 2015), which may lead to a reduced association between motivation measures that assess different domains. Further, it may be that over time, a person's perception of their motivation becomes less aligned with behavior. For example, a mismatch between perception and behavior might result from prolonged experience with factors such as stigma (Firmin, Luther, Lysaker, Minor, & Salyers, 2016), defeatist performance beliefs (Grant & Beck, 2009), less stimulating environments (e.g., a hospital), and unfulfilled basic psychological needs (Ryan & Deci, 2000b). Future research is needed to clarify factors or additional moderators that might impact the relationship between self-reported and clinician-rated motivation measures, especially as the length of illness increases.

In part to identify whether individual motivation domains would also help to explain the observed heterogeneity, we explored the relationships across assessment types within domains for intrinsic and amotivation domains. No studies assessed the relationship between different extrinsic motivation measures types, so this relationship was unable to be analyzed. Contrary to our hypothesis, the overall relationship between self-reported and clinician-rated intrinsic motivation was small and non-significant. However, few studies were included, and the self-reported intrinsic motivation measures largely assessed intrinsic motivation for a specific task, while the clinician-rated measure assessed intrinsic motivation for tasks more generally. As J. Choi et al. (2014) suggested, the limited relationship may in part be due to the fact that

motivation for different tasks was assessed, but additional research is needed to clarify this relationship. Further, in line with our hypothesis, the overall mean effect size between self-reported and clinician-rated amotivation measures was positive, significant, and represented a medium effect size. Notably, this overall effect size was relatively larger than the overall effect size between all included self-reported and clinician-rated motivation measures, suggesting that amotivation measures across measurement types may be more strongly related than measures assessing different motivation domains (i.e., relationship between self-reported intrinsic motivation and clinician-rated amotivation). Indeed, this is consistent with SDT, which describes these motivation domains as distinct yet possibly co-occurring constructs (Ryan & Deci, 2000a). For instance, it is possible that someone is motivated to do something both because of the pleasure one derives from the activity as well as the monetary reward associated with it (e.g., writing and publishing books). At other times, people may be extrinsically motivated but not intrinsically motivated to complete a specific task. Therefore, we may not always expect these different types of motivation to strongly correlate with one another. However, given that we were unable to conduct one of the domain analyses and few studies were included in the intrinsic motivation domain meta-analyses, additional work clarifying the overlap both within and between different motivation domains is needed. Further, our results may be impacted by how we categorized measures into different motivation domains. Indeed, Deci and Ryan's (2008) more recent macrotheory of motivation differentiates motivation domains into autonomous motivation, controlled motivation, and amotivation. Notably, this newer theory describes autonomous and controlled motivation as involving both internal and external processes. Thus, future research may benefit from exploring the association of these more recent motivation domains as well as clarifying which measures assess these constructs. Similarly, although Deci

and Ryan's (1985a) seminal work on differentiating motivation into the domains of intrinsic motivation, extrinsic motivation, and amotivation has served as a critical guide for schizophrenia motivation research, future work focused on the conceptualization and measurement of motivation in schizophrenia may benefit from incorporating Deci and Ryan's (2008) more recent theory.

It is also useful to frame these results in the context of different psychometric theories and methods. Specifically, examining the overlap between different methods of assessing motivation is consistent with the Multitrait-Multimethod Matrix (MTMM; Campbell & Fiske, 1959), which is a statistical approach that aims to examine the suitability of tests to measure a specified construct (i.e., construct validity). Briefly, Campbell and Fiske (1959) described that in order for tests to have adequate construct validity, they must demonstrate adequate convergent and discriminant validity with different measurement methods of the same construct (i.e., self-reported, clinician-rated motivation measures) and distinct traits that are not theoretically related (i.e., positive symptoms), respectively. Further, they, along with others (Pitoniak, Sireci, & Luecht, 2002), argued that in order for adequate convergent validity to be established, correlations among the different measurement methods of a construct should be significantly different from zero and be "sufficiently large to encourage further examination of validity" (Campbell & Fiske, 1959, p. 83). Although additional research on the discriminant validity of the included motivation measures is needed to adequately use the MTMM, given that the largest overall effect size observed between performance-based and other methods of motivation assessment was only small in magnitude, it is uncertain if performance-based motivation measures meet MTMM's criteria for adequate convergent and thus construct validity. However, some have noted (Cohen, 2016) that standard psychometric methods may fail to accurately

evaluate the complex associations with more novel objective approaches such as performance-based measures of motivation. Instead, researchers suggest that methods such as argument-based validity (see Kane, 1992 for more information on this approach) might be a better means to assess the validity of these measures. Despite these different approaches to examining the validity of measures, it is clear that additional research examining the construct validity of performance-based as well as self-reported and clinician-rated motivation measures is needed in order to bolster our assumptions that these measures are truly assessing motivation.

In addition to other limitations common to all meta-analyses (e.g., limited by the constraints of the primary studies; Borenstein et al., 2009), there are also several study-specific limitations that should be considered when interpreting the findings from this study. First, we included only studies that were available in English. Second, despite a large number of studies identified during the initial literature search, few were able to be included in the current study (in part due to the recency of some assessment methods), which limited our ability to conduct all proposed meta-analyses and moderator analyses. However, many of the included study-level correlations were not previously published and were obtained through contacting researchers, which reduces the possibility of publication bias. In addition, although we used SDT and recent literature reviews on motivation and negative symptom measures to guide our inclusion of measures, we did not include all possible measures of “motivation.” This was partially in attempt to reduce construct validity concerns and to address one of the common criticisms of meta-analyses: comparing apples to oranges (Borenstein et al., 2009), or in this case, comparing motivation measures that might be assessing constructs that are too disparate from one another. However, because our goal was to draw conclusions about the way the literature is currently assessing a particular “fruit” (i.e., motivation), we still aimed to include a range of measures that



have been used to measure motivation in the schizophrenia literature by basing our measure inclusion on multiple reviews that describe a large array of the most widely-used motivation and negative symptom measures. In addition, as previously discussed, many scores can be derived from performance-based measures, and there is no completely agreed upon score that is used to assess “motivation.” To reduce issues surrounding interpretation of different scoring methods (Reddy, Horan, & Green, 2015), when appropriate, this study used the most commonly used motivation score of percent of hard choices chosen across trials. However, future work is needed to identify if other scoring methods (e.g. creating a difference score across different conditions; Horan et al., 2015) might be more precise measures of “motivation” or at least more in line with self-reported and clinician-rated motivation.

Despite these limitations, our findings point to several areas of research related to the assessment of motivation in schizophrenia. Most importantly, additional work examining the construct validity (i.e., convergent and discriminant validity), particularly of performance-based motivation measures, is needed. Ideally, during these validation efforts, researchers will also utilize more rigorous methods, such as the MTMM or argument-based validity, to examine the construct validity of these motivation scales. Second, future research clarifying the construct or convergent validity of motivation scales may also benefit by further investigating their relationship with overall negative symptoms scores. This is particularly relevant given that several of the first-generation negative symptoms measures such as the SANS or the PANSS have been criticized for not adequately assessing internal experiences or the full range of each negative symptom domain (Blanchard et al., 2010), while the second-generation negative symptom measures such as the CAINS and the BNSS have been developed in part to address these limitations of the first-generation negative symptom assessments. Similarly, as noted

above, researchers have identified that there are two key subdomains of negative symptoms (Green et al., 2012; Strauss et al., 2013), with amotivation or the experiential negative symptoms being one of the domains. However, depending on the scale or factor analyses (Blanchard & Cohen, 2005; Foussias & Remington, 2010; Liemburg et al., 2013), anhedonia is at times included in the experiential negative symptom subscale. Although amotivation and anhedonia are thought of as distinct (yet partly overlapping) negative symptoms (Foussias & Remington, 2010), future research could also benefit from clarifying the overlap between these symptoms as well as examining whether there is greater concordance between measurement types when looking at both amotivation and anhedonia. An additional important area of future research involves identifying what are the “best” or “gold-standard” measures of motivation across each measurement type, or similarly, what measures or combinations of measures most closely resemble the different motivation domains and the multidimensional construct of “motivation.” Finally, as others have suggested (J. Choi et al., 2014; Fervaha, Foussias, et al., 2015), another important area of future research is to examine what type of motivation measurement (as well as motivation domains) may be most predictive of functioning and other symptoms.

The findings from these meta-analyses can also be used to guide and improve our assessment of motivation in schizophrenia. These findings suggest that these three motivation measurement types are at best only assessing a small amount of a shared underlying construct. Therefore, in contrast to some current practices, these three motivation measurement types should not be used interchangeably to assess “motivation.” Instead, at the outset, researchers and clinicians should refer to each measure as the specific construct or domain it was designed to measure or capture. Indeed, recent work on motivation in schizophrenia (J. Choi et al., 2014; K.-H. Choi, Saperstein, & Medalia, 2012; Luther et al., 2015) has begun to use more precise

language when describing motivation measures. Although more work clarifying the construct validity of these scales is needed, future researchers may also follow this trend and describe the included clinician-rated scales as scales assessing more trait levels of general intrinsic motivation (or amotivation) or describe performance-based measures as measures assessing state levels of extrinsic motivation for monetary rewards or state levels of effort-based decision-making for monetary rewards. Similarly, self-reported measures of motivation should be accurately described. For example, these measures could be described as assessing state motivation for a specific task or as assessing trait-like motivation for a range of activities or tasks instead of measures of “motivation.” Taken together, our findings highlight the limited overlap between motivation measures and suggest that additional research involving the assessment of motivation is needed to help improve our understanding and treatment of motivation in people with schizophrenia.

## References

- Andreasen, N. C. (1984). *Scale for the assessment of negative symptoms (SANS)*: Department of Psychiatry, College of Medicine, The University of Iowa.
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5<sup>th</sup> ed.). Arlington, VA: American Psychiatric Publishing.
- Balshem, H., Stevens, A., Ansari, M., Norris, S., Kansagara, D., Shamliyan, T., Chou, R., Chung, M., Moher, D., Dickersin, K. (2013). Finding grey literature evidence and assessing for outcome and analysis reporting biases when comparing medical interventions: AHRQ and the effective health care program. (AHRQ Publication No. 13(14)-EHC096-EF). Rockville, MD: Agency for Healthcare Research and Quality.
- Barch, D. M., Gold, J. M., & Kring, A. M. (2017). Paradigms for assessing hedonic processing and motivation in humans: relevance to understanding negative symptoms in psychopathology. *Schizophrenia Bulletin*, 43(4), 701-705.
- Barch, D. M., Yodkovik, N., Sypher-Locke, H., & Hanewinkel, M. (2008). Intrinsic motivation in schizophrenia: relationships to cognitive function, depression, anxiety, and personality. *Journal of Abnormal Psychology*, 117(4), 776-787.
- Blanchard, J. J., & Cohen, A. S. (2005). The structure of negative symptoms within schizophrenia: implications for assessment. *Schizophrenia Bulletin*, 32(2), 238-245.
- Blanchard, J. J., Kring, A. M., Horan, W. P., & Gur, R. (2010). Toward the next generation of negative symptom assessments: The collaboration to advance negative symptom assessment in schizophrenia. *Schizophrenia Bulletin*, 37(2): 291–299.
- Bleuler, E. (1950). *Dementia Praecox or the group of schizophrenias*. (J. Zinkin, Trans.). New York, NY: International Universities Press. (Original work published 1911)

- Borenstein, M., Hedges, L. V., Higgins, J., & Rothstein, H. R. (2009). *Introduction to Meta-analysis*. Chichester, UK: John Wiley & Sons, Ltd.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2014). *Comprehensive Meta-Analysis (Version 3)*. Englewood, NJ: Biostat.
- Breitborde, N. J., Woolverton, C., Frost, R. B., & Kiewel, N. A. (2014). Self-determination theory and first-episode psychosis: a replication. *Advances in Psychiatry*, 2014.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81-105.
- Card, N. A. (2012). *Applied meta-analysis for social science research*. New York, NY: Guilford Press.
- Carter, C. S., Barch, D. M., Buchanan, R. W., Bullmore, E., Krystal, J. H., Cohen, J., . . . Robbins, T. (2008). Identifying cognitive mechanisms targeted for treatment development in schizophrenia: An overview of the first meeting of the Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia Initiative. *Biological Psychiatry*, 64(1), 4-10.
- Choi, J., Choi, K.-H., Reddy, L. F., & Fiszdon, J. M. (2014). Measuring motivation in schizophrenia: Is a general state of motivation necessary for task-specific motivation? *Schizophrenia Research*, 153(1), 209-213.
- Choi, J., Fiszdon, J. M., & Medalia, A. (2010). Expectancy-value theory in persistence of learning effects in schizophrenia: Role of task value and perceived competency. *Schizophrenia Bulletin*, 36(5), 957-965.
- Choi, J., Mogami, T., & Medalia, A. (2010). Intrinsic motivation inventory: An adapted measure for schizophrenia research. *Schizophrenia Bulletin*, 36(5): 966-976.

- Choi, K.-H., Saperstein, A. M., & Medalia, A. (2012). The relationship of trait to state motivation: The role of self-competency beliefs. *Schizophrenia Research*, 139(1), 73-77.
- Cohen, A.S. (September 2016). Psychometrics for contemporary psychopathology assessment. Paper presented at the Thirtieth Annual Meeting of the Society for Research in Psychopathology, Baltimore, MD.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.
- Cooper, S., Lavaysse, L. M., & Gard, D. E. (2015). Assessing motivation orientations in schizophrenia: Scale development and validation. *Psychiatry Research*, 225(1), 70-78.
- Deci, E. L., & Ryan, R. M. (1985a). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum Press.
- Deci, E. L., & Ryan, R. M. (1985b). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality*, 19(2), 109-134.
- Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. *Canadian Psychology/Psychologie Canadienne*, 49(3), 182-185.
- DeRosse, P., Nitzburg, G. C., Kompancaril, B., & Malhotra, A. K. (2014). The relation between childhood maltreatment and psychosis in patients with schizophrenia and non-psychiatric controls. *Schizophrenia Research*, 155(1), 66-71.
- Eden, J., Levit, L., Berg, A., & Morton, S. (Eds). (2011). *Finding what works in health care: Standards for systematic reviews*. Washington, DC: National Academies Press.
- Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109), 629-634.
- Engel, M., & Lincoln, T. M. (2016). Motivation and Pleasure Scale-Self-Report (MAP-SR): Validation of the German version of a self-report measure for screening negative symptoms in schizophrenia. *Comprehensive Psychiatry*, 65, 110-115.

- Evensen, J., Røssberg, J. I., Barder, H., Haahr, U., ten Velden Hegelstad, W., Joa, I., ... & Rund, B. R. (2012). Apathy in first episode psychosis patients: A ten year longitudinal follow-up study. *Schizophrenia Research*, 136(1), 19-24.
- Faerden, A., Friis, S., Agartz, I., Barrett, E. A., Nesvåg, R., Finset, A., & Melle, I. (2009). Apathy and functioning in first-episode psychosis. *Psychiatric Services*, 60(11), 1495-1503.
- Faerden, A., Nesvåg, R., Barrett, E. A., Agartz, I., Finset, A., Friis, S., ... Melle, I. (2008). Assessing apathy: The use of the Apathy Evaluation Scale in first episode psychosis. *European Psychiatry*, 23(1), 33-39.
- Fervaha, G., Foussias, G., Agid, O., & Remington, G. (2014). Motivational and neurocognitive deficits are central to the prediction of longitudinal functional outcome in schizophrenia. *Acta Psychiatrica Scandinavica*, 130(4), 290-299.
- Fervaha, G., Duncan, M., Foussias, G., Agid, O., Faulkner, G. E., & Remington, G. (2015). Effort-based decision making as an objective paradigm for the assessment of motivational deficits in schizophrenia. *Schizophrenia Research*, 168(1), 483-490.
- Fervaha, G., Foussias, G., Takeuchi, H., Agid, O., & Remington, G. (2015). Measuring motivation in people with schizophrenia. *Schizophrenia Research*, 169(1), 423-426.
- Fervaha, G., Graff-Guerrero, A., Zakzanis, K. K., Foussias, G., Agid, O., & Remington, G. (2013). Incentive motivation deficits in schizophrenia reflect effort computation impairments during cost-benefit decision-making. *Journal of Psychiatric Research*, 47(11), 1590-1596.

Fervaha, G., Zakzanis, K. K., Foussias, G., Graff-Guerrero, A., Agid, O., & Remington, G.

(2014). Motivational deficits and cognitive test performance in schizophrenia. *JAMA*

*Psychiatry*, 71(9), 1058-1065.

Firmin, R. L., Luther, L., Lysaker, P. H., Minor, K. S., & Salyers, M. P. (2016). Stigma resistance is associated with positive psychiatric and psychosocial outcomes: A meta- analysis. *Schizophrenia Research*, 175, 118-128.

Foussias, G., Mann, S., Zakzanis, K., van Reekum, R., Agid, O., & Remington, G. (2011).

Prediction of longitudinal functional outcomes in schizophrenia: the impact of baseline motivational deficits. *Schizophrenia Research*, 132(1), 24-27.

Foussias, G., & Remington, G. (2010). Negative symptoms in schizophrenia: Avolition and Occam's razor. *Schizophrenia Bulletin*, 36(2), 359-369.

Fu, R., Gartlehner, G., Grant, M., Shamliyan, T., Sedrakyan, A., Wilt, T. J., . . . Ismaila, A.

(2011). Conducting quantitative synthesis when comparing medical interventions: AHRQ and the Effective Health Care Program. *Journal of Clinical Epidemiology*, 64(11), 1187-1197.

Gard, D. E., Sanchez, A. H., Starr, J., Cooper, S., Fisher, M., Rowlands, A., & Vinogradov, S.

(2014). Using self-determination theory to understand motivation deficits in schizophrenia: The 'why' of motivated behavior. *Schizophrenia Research*, 156(2), 217-222.

Gillan, C. M., Papmeyer, M., Morein-Zamir, S., Sahakian, B. J., Fineberg, N. A., Robbins, T.

W., & de Wit, S. (2011). Disruption in the balance between goal-directed behavior and habit learning in obsessive-compulsive disorder. *American Journal of Psychiatry*, 168(7), 718-726.



- Glynn, S., & Mueser, K. T. (1986). Social learning for chronic mental inpatients. *Schizophrenia Bulletin*, 12(4), 648-668.
- Gold, J. M., Strauss, G. P., Waltz, J. A., Robinson, B. M., Brown, J. K., & Frank, M. J. (2013). Negative symptoms of schizophrenia are associated with abnormal effort-cost computations. *Biological Psychiatry*, 74(2), 130-136.
- Gold, J. M., Waltz, J. A., Prentice, K. J., Morris, S. E., & Heerey, E. A. (2008). Reward processing in schizophrenia: A deficit in the representation of value. *Schizophrenia Bulletin*, 34(5), 835-847.
- Grant, P. M., & Beck, A. T. (2008). Defeatist beliefs as a mediator of cognitive impairment, negative symptoms, and functioning in schizophrenia. *Schizophrenia Bulletin*, 35(4), 798-806.
- Green, M. F., Helleman, G., Horan, W. P., Lee, J., & Wynn, J. K. (2012). From perception to functional outcome in schizophrenia: Modeling the role of ability and motivation. *Archives of General Psychiatry*, 69(12), 1216-1224.
- Green, M. F., & Horan, W. P. (2015). Effort-based decision making in schizophrenia: Evaluation of paradigms to measure motivational deficits. *Schizophrenia Bulletin*, 41(5), 1021-1023.
- Green, M. F., Horan, W. P., Barch, D. M., & Gold, J. M. (2015). Effort-based decision making: A novel approach for assessing motivation in schizophrenia. *Schizophrenia Bulletin*, 41(5), 1035-1044.
- Gupta, M., Holshausen, K., Gou, L., & Bowie, C. (2014). Measuring negative symptom change in schizophrenia: Considering alternatives to self-report. *Expert Review of Neurotherapeutics*, 14(8), 911-922.

- Heerey, E. A., Robinson, B. M., McMahon, R. P., & Gold, J. M. (2007). Delay discounting in schizophrenia. *Cognitive Neuropsychiatry*, *12*(3), 213-221.
- Heinrichs, D. W., Hanlon, T. E., & Carpenter, W. T. (1984). The Quality of Life Scale: An instrument for rating the schizophrenic deficit syndrome. *Schizophrenia Bulletin*, *10*(3), 388-398.
- Higgins, J.P.T., & Green, S. (Eds.) (2011). *Cochrane Handbook for Systematic Reviews of Interventions* (Version 5.1.0). The Cochrane Collaboration. Available from <http://handbook.cochrane.org>.
- Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539-1558.
- Horan, W. P., Kring, A. M., Gur, R. E., Reise, S. P., & Blanchard, J. J. (2011). Development and psychometric validation of the Clinical Assessment Interview for Negative Symptoms (CAINS). *Schizophrenia Research*, *132*(2), 140-145.
- Horan, W. P., Reddy, L. F., Barch, D. M., Buchanan, R. W., Dunayevich, E., Gold, J. M., . . . Green, M. F. (2015). Effort-based decision-making paradigms for clinical trials in schizophrenia: Part 2—external validity and correlates. *Schizophrenia Bulletin*, *41*(5), 1055-1065.
- Huedo-Medina, T. B., Sánchez-Meca, J., Marín-Martínez, F., & Botella, J. (2006). Assessing heterogeneity in meta-analysis: Q statistic or  $I^2$  index? *Psychological Methods*, *11*(2), 193-206.
- Insel, C., Reinen, J., Weber, J., Wager, T. D., Jarskog, L. F., Shohamy, D., & Smith, E. E. (2014). Antipsychotic dose modulates behavioral and neural responses to feedback

- during reinforcement learning in schizophrenia. *Cognitive, Affective, & Behavioral Neuroscience*, 14(1), 189-201.
- Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112(3), 527-535.
- Kay, S. R., Fiszbein, A., & Opfer, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, 13(2), 261-276.
- Kepes, S., Banks, G. C., McDaniel, M., & Whetzel, D. L. (2012). Publication bias in the organizational sciences. *Organizational Research Methods*, 15(4), 624-662.
- Kiang, M., Christensen, B. K., Remington, G., & Kapur, S. (2003). Apathy in schizophrenia: Clinical correlates and association with functional outcome. *Schizophrenia Research*, 63(1), 79-88.
- Kirkpatrick, B., Fenton, W. S., Carpenter, W. T., & Marder, S. R. (2006). The NIMH-MATRICES Consensus Statement on Negative Symptoms. *Schizophrenia Bulletin*, 32(2), 214-219.
- Kirkpatrick, B., Strauss, G. P., Nguyen, L., Fischer, B. A., Daniel, D. G., Cienfuegos, A., & Marder, S. R. (2011). The Brief Negative Symptom Scale: Psychometric properties. *Schizophrenia Bulletin*, 37(2), 300-305.
- Kirsch, P., Ronshausen, S., Mier, D., & Gallhofer, B. (2007). The influence of antipsychotic treatment on brain reward system reactivity in schizophrenia patients. *Pharmacopsychiatry*, 40(5), 196-198.
- Kleinginna, P. R., & Kleinginna, A. M. (1981) A categorized list of motivation definitions, with a suggestion for a consensual definition. *Motivation and Emotion*, 5(3), 263-291.
- Kraepelin, E. (1919). Dementia praecox and paraphrenia. (M. Barclay. Trans.). Chicago, IL: Chicago Medical Book Company. (Original work published 1913)

- Kremen, L. C., Fiszdon, J. M., Kurtz, M. M., Silverstein, S. M., & Choi, J. (2016). Intrinsic and extrinsic motivation and learning in schizophrenia. *Current Behavioral Neuroscience Reports*, 3(2), 144-153.
- Kring, A. M., Gur, R. E., Blanchard, J. J., Horan, W. P., & Reise, S. P. (2013). The Clinical Assessment Interview for Negative Symptoms (CAINS): Final development and validation. *American Journal of Psychiatry*, 170(2), 165-172.
- Liddle, P. F., & Barnes, T. R. (1988). The subjective experience of deficits in schizophrenia. *Comprehensive Psychiatry*, 29(2), 157-164.
- Liemburg, E., Castelein, S., Stewart, R., van der Gaag, M., Aleman, A., Knegtering, H., & Risk, G. (2013). Two subdomains of negative symptoms in psychotic disorders: Established and confirmed in two large cohorts. *Journal of Psychiatric Research*, 47(6), 718-725.
- Lincoln, T. M., Dollfus, S., & Lyne, J. (2017). Current developments and challenges in the assessment of negative symptoms. *Schizophrenia Research*, 186, 8-18.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: SAGE Publications, Inc.
- Llerena, K., Park, S. G., McCarthy, J. M., Couture, S. M., Bennett, M. E., & Blanchard, J. J. (2013). The Motivation and Pleasure Scale–Self-Report (MAP-SR): Reliability and validity of a self-report measure of negative symptoms. *Comprehensive Psychiatry*, 54(5), 568-574.
- Luther, L., Firmin, R. L., Minor, K. S., Vohs, J. L., Buck, B., Buck, K. D., & Lysaker, P. H. (2016). Metacognition deficits as a risk factor for prospective motivation deficits in schizophrenia spectrum disorders. *Psychiatry Research*, 245, 172-178.

- Luther, L., Firmin, R. L., Vohs, J. L., Buck, K. D., Rand, K. L., & Lysaker, P. H. (2016). Intrinsic motivation as a mediator between metacognition deficits and impaired functioning in psychosis. *British Journal of Clinical Psychology*, 55(3), 332-47.
- Luther, L., Lysaker, P. H., Firmin, R. L., Breier, A., & Vohs, J. L. (2015). Intrinsic motivation and amotivation in first episode and prolonged psychosis. *Schizophrenia Research*, 169(1), 418-422.
- Marder, S. R., & Galderisi, S. (2017). The current conceptualization of negative symptoms in schizophrenia. *World Psychiatry*, 16(1), 14-24.
- Marin, R. S., Biedrzycki, R. C., & Firinciogullari, S. (1991). Reliability and validity of the Apathy Evaluation Scale. *Psychiatry Research*, 38(2), 143-162.
- Markou, A., Salamone, J. D., Bussey, T. J., Mar, A. C., Brunner, D., Gilmour, G., & Balsam, P. (2013). Measuring reinforcement learning and motivation constructs in experimental animals: Relevance to the negative symptoms of schizophrenia. *Neuroscience & Biobehavioral Reviews*, 37(9), 2149-2165.
- McCarthy, J. M., Treadway, M. T., Bennett, M. E., & Blanchard, J. J. (2016). Inefficient effort allocation and negative symptoms in individuals with schizophrenia. *Schizophrenia Research*, 170(2), 278-284.
- Medalia, A., & Brekke, J. (2010). In search of a theoretical structure for understanding motivation in schizophrenia. *Schizophrenia Bulletin*, 36(5): 912-918.
- Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine*, 151(4), 264-269.

- Nakagami, E., Xie, B., Hoe, M., & Brekke, J. S. (2008). Intrinsic motivation, neurocognition and psychosocial functioning in schizophrenia: Testing mediator and moderator effects. *Schizophrenia Research*, 105(1), 95-104.
- O'Driscoll, C., Laing, J., & Mason, O. (2014). Cognitive emotion regulation strategies, alexithymia and dissociation in schizophrenia, a review and meta- analysis. *Clinical Psychology Review*, 34(6), 482-495.
- Pearlman, D. M., & Najjar, S. (2014). Meta-analysis of the association between N- methyl-d- aspartate receptor antibodies and schizophrenia, schizoaffective disorder, bipolar disorder, and major depressive disorder. *Schizophrenia Research*, 157(1), 249-258.
- Pitoniak, M. J., Sireci, S. G., & Luecht, R. M. (2002). A multitrait-multimethod validity investigation of scores from a professional licensure examination. *Educational and Psychological Measurement*, 62(3), 498-516.
- Reddy, L. F., Horan, W. P., Barch, D. M., Buchanan, R. W., Dunayevich, E., Gold, J. M., . . . Green, M. F. (2015). Effort-Based Decision-Making Paradigms for Clinical Trials in Schizophrenia: Part 1-Psychometric Characteristics of 5 Paradigms. *Schizophrenia Bulletin*, 41(5), 1045-1054.
- Reddy, L. F., Horan, W. P., & Green, M. F. (2015). Motivational deficits and negative symptoms in schizophrenia: Concepts and assessments. In E. H. Simpson & P.D. Balsam (Eds.), *Behavioral Neuroscience of Motivation* (357-373). Switzerland: Springer International Publishing.
- Ryan, R. M., & Deci, E. L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54-67.

- Ryan, R. M., & Deci, E.L. (2000b). Self-Determination Theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), 68-78.
- Salyers, M. P., Bonfils, K. A., Luther, L., Firmin, R. L., White, D. A., Adams, E. L., & Rollins, A. L. (2017). The relationship between professional burnout and quality and safety in healthcare: A meta-analysis. *Journal of General Internal Medicine*, 32(4), 475-482.
- Saperstein, A. M., Fiszdon, J. M., & Bell, M. D. (2011). Intrinsic motivation as a predictor of work outcome after vocational rehabilitation in schizophrenia. *Journal of Nervous and Mental Disease*, 199(9), 672-677.
- Selten, J.-P. C., Sijben, N. E., Van den Bosch, R. J., Omloo-Visser, J., & Warmerdam, H. (1993). The subjective experience of negative symptoms: A self-rating scale. *Comprehensive Psychiatry*, 34(3), 192-197.
- Strauss, G. P., Frank, M. J., Waltz, J. A., Kasanova, Z., Herbener, E. S., & Gold, J. M. (2011). Deficits in positive reinforcement learning and uncertainty-driven exploration are associated with distinct aspects of negative symptoms in schizophrenia. *Biological Psychiatry*, 69(5), 424-431.
- Strauss, G. P., Horan, W. P., Kirkpatrick, B., Fischer, B. A., Keller, W. R., Miski, P., . . . Carpenter, W. T. (2013). Deconstructing negative symptoms of schizophrenia: avolition–apathy and diminished expression clusters predict clinical presentation and functional outcome. *Journal of Psychiatric Research*, 47(6), 783-790.
- Strauss, G. P., Keller, W. R., Buchanan, R. W., Gold, J. M., Fischer, B. A., McMahon, R. P., . . . Kirkpatrick, B. (2012). Next-generation negative symptom assessment for clinical trials: Validation of the Brief Negative Symptom Scale. *Schizophrenia Research*, 142(1), 88-92.

- Strauss, G. P., Waltz, J. A., & Gold, J. M. (2014). A review of reward processing and motivational impairment in schizophrenia. *Schizophrenia Bulletin*, 40(Suppl 2), S107-S116.
- Strauss, G. P., Whearty, K. M., Morra, L. F., Sullivan, S. K., Ossenfort, K. L., & Frost, K. H. (2016). Avolition in schizophrenia is associated with reduced willingness to expend effort for reward on a progressive ratio task. *Schizophrenia Research*, 170(1), 198-204.
- Takeuchi, H., Fervaha, G., & Remington, G. (2016). Reliability of a patient-reported outcome measure in schizophrenia: Results from back-to-back self-ratings. *Psychiatry Research*, 244, 415-419.
- Treadway, M. T., Buckholz, J. W., Schwartzman, A. N., Lambert, W. E., & Zald, D. H. (2009). Worth the 'EEfRT'? The Effort Expenditure for Rewards Task as an objective measure of motivation and anhedonia. *PLoS ONE*, 4(8), e6598.
- Treadway, M. T., Peterman, J. S., Zald, D. H., & Park, S. (2015). Impaired effort allocation in patients with schizophrenia. *Schizophrenia Research*, 161(2), 382-385.
- Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How many studies do you need? A primer on statistical power for meta-analysis. *Journal of Educational and Behavioral Statistics*, 35(2), 215-247.
- Werry, J. S. (1969). The token economy: A motivational system for therapy and rehabilitation. *Archives of General Psychiatry*, 21(1), 121-122.
- Yamada, A.-M., Lee, K. K., Dinh, T. Q., Barrio, C., & Brekke, J. S. (2010). Intrinsic motivation as a mediator of relationships between symptoms and functioning among individuals with schizophrenia spectrum disorders in a diverse urban community. *The Journal of Nervous and Mental Disease*, 198(1), 28-34.



### Role of Funding Sources

This publication was supported by a Pre-doctoral Award from the Indiana Clinical and Translational Sciences Institute (I-CTSI), which is funded in part by a National Institutes of Health (NIH), National Center for Advancing Translational Sciences, Clinical and Translational Sciences Award (# UL1TR001108). The I-CTSI and NIH had no role in the study design, collection, analysis or interpretation of the data, manuscript preparation, or the decision to submit the paper for publication.

### Contributors

Author Luther, L. conducted the literature search, performed the data analysis, wrote the first draft of the manuscript, and finalized the manuscript. Authors Firmin, R.L., Lysaker, P.H., Minor, K.S., and Salyers, M.P., provided conceptual guidance in the design of the literature review, analyses, and manuscript and assisted with revising the manuscript. All authors approved the final version of the manuscript.

### Conflicts of Interest

The authors declare that they have no conflicts of interest.

### Acknowledgements

The authors are incredibly grateful for the researchers who provided additional data for this meta-analysis. Some preliminary analyses were presented at the 30<sup>th</sup> annual meeting for the Society for Research on Psychopathology and the 8<sup>th</sup> annual meeting of the Indiana Clinical and Translational Sciences Institute.

Figure 1. PRISMA Study Retrieval Flow Diagram

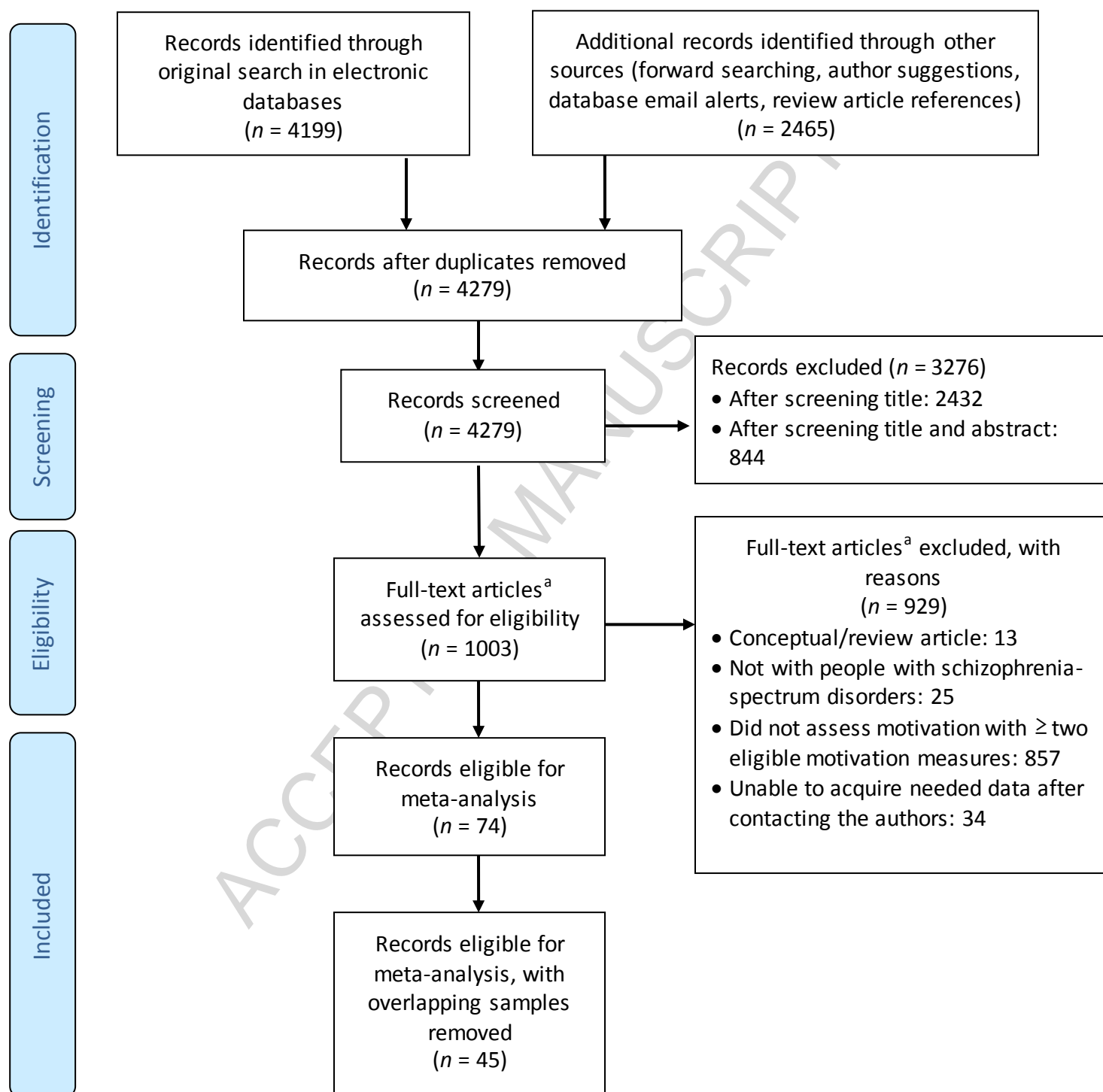
<sup>a</sup>Includes conference abstracts

Figure 2. Forest plot of the meta-analysis examining the relationship between self-reported and clinician-rated motivation measures ( $k = 33$ )

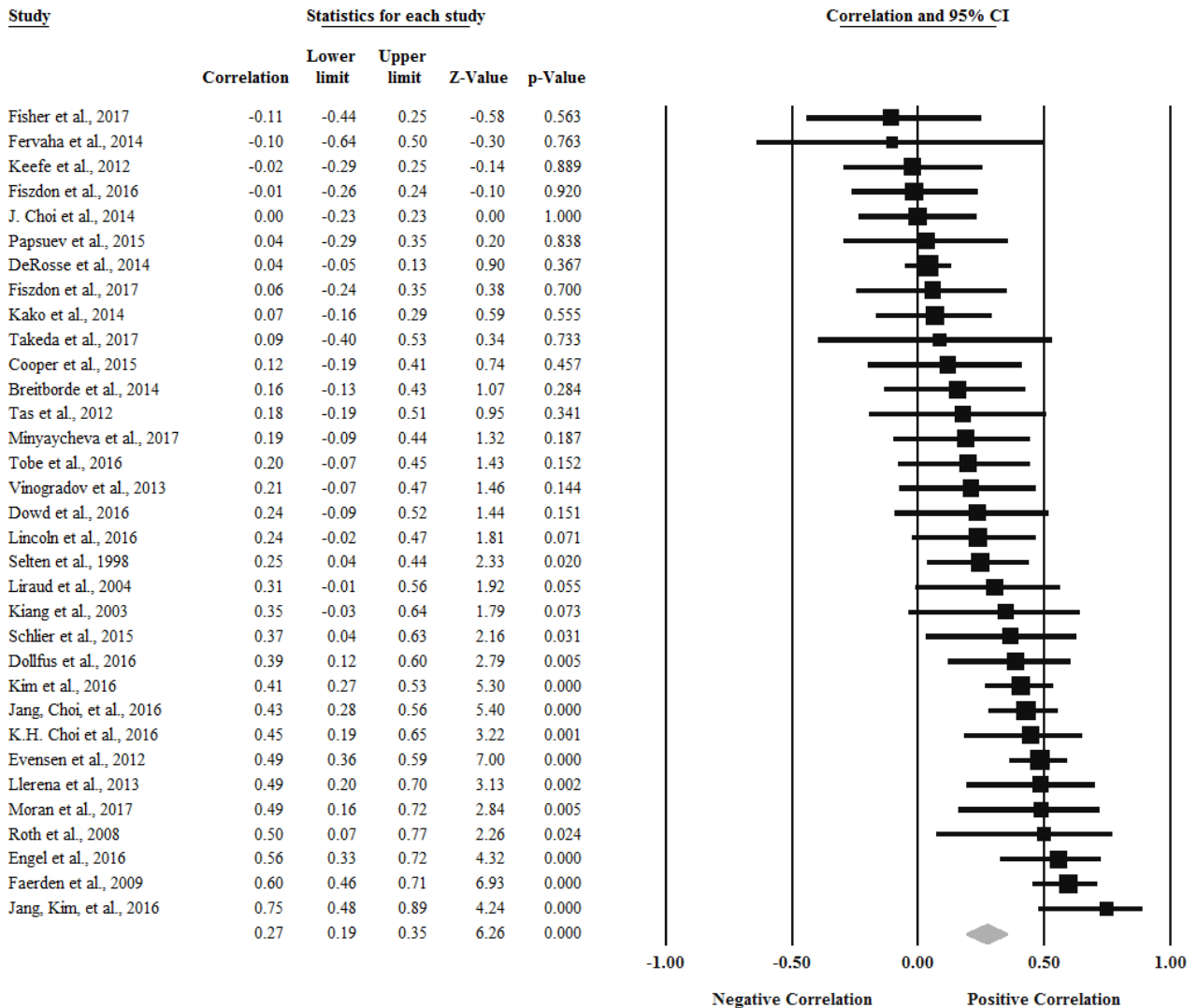


Figure 3. Forest plot of the meta-analysis examining the relationship between performance-based and clinician-rated motivation measures ( $k = 11$ )

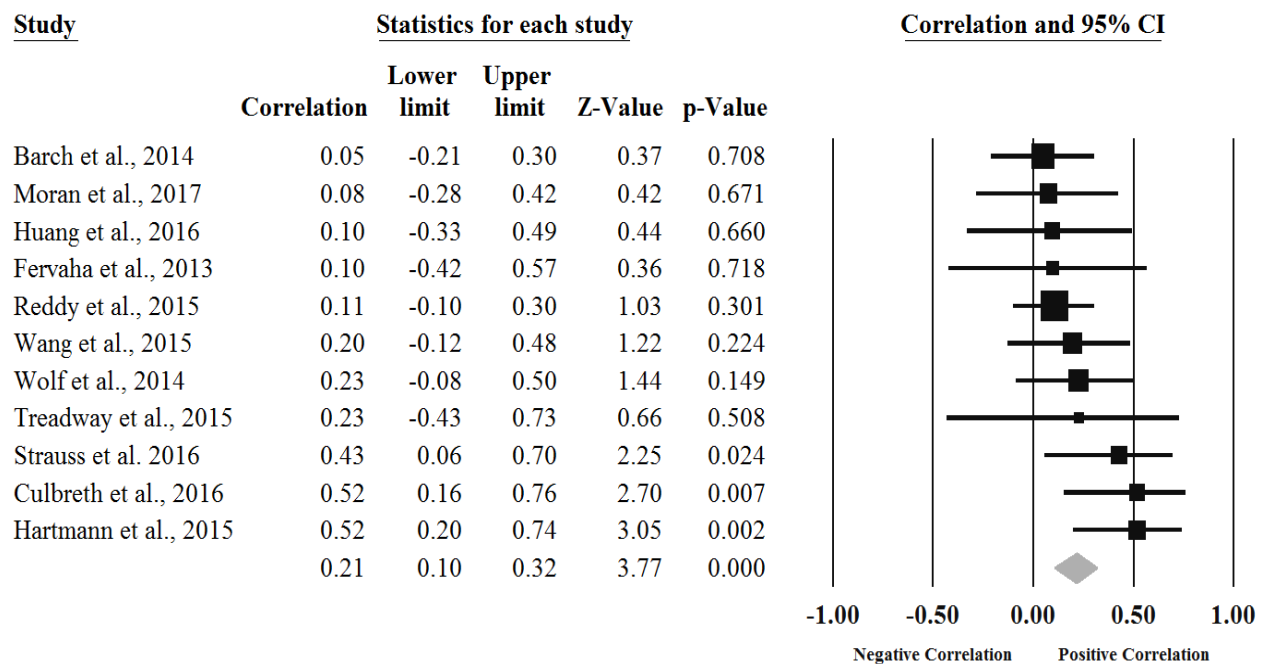


Figure 4. Forest plot of the meta-analysis examining the relationship between self-reported and performance-based motivation measures ( $k = 2$ )

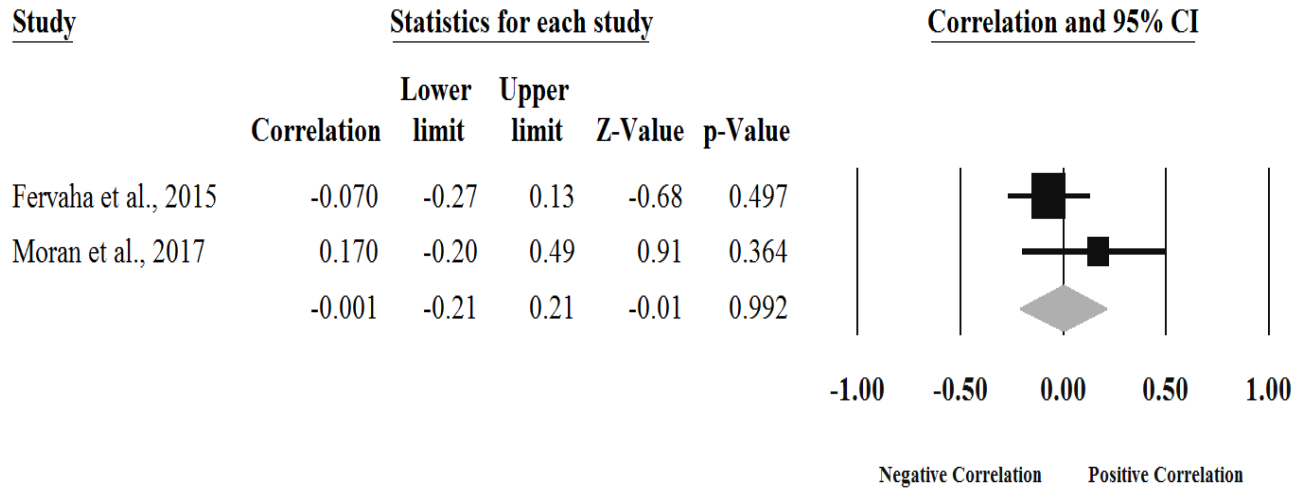


Figure 5. Forest plot of the meta-analysis examining the relationship between self-reported and clinician-rated amotivation measures ( $k = 23$ )

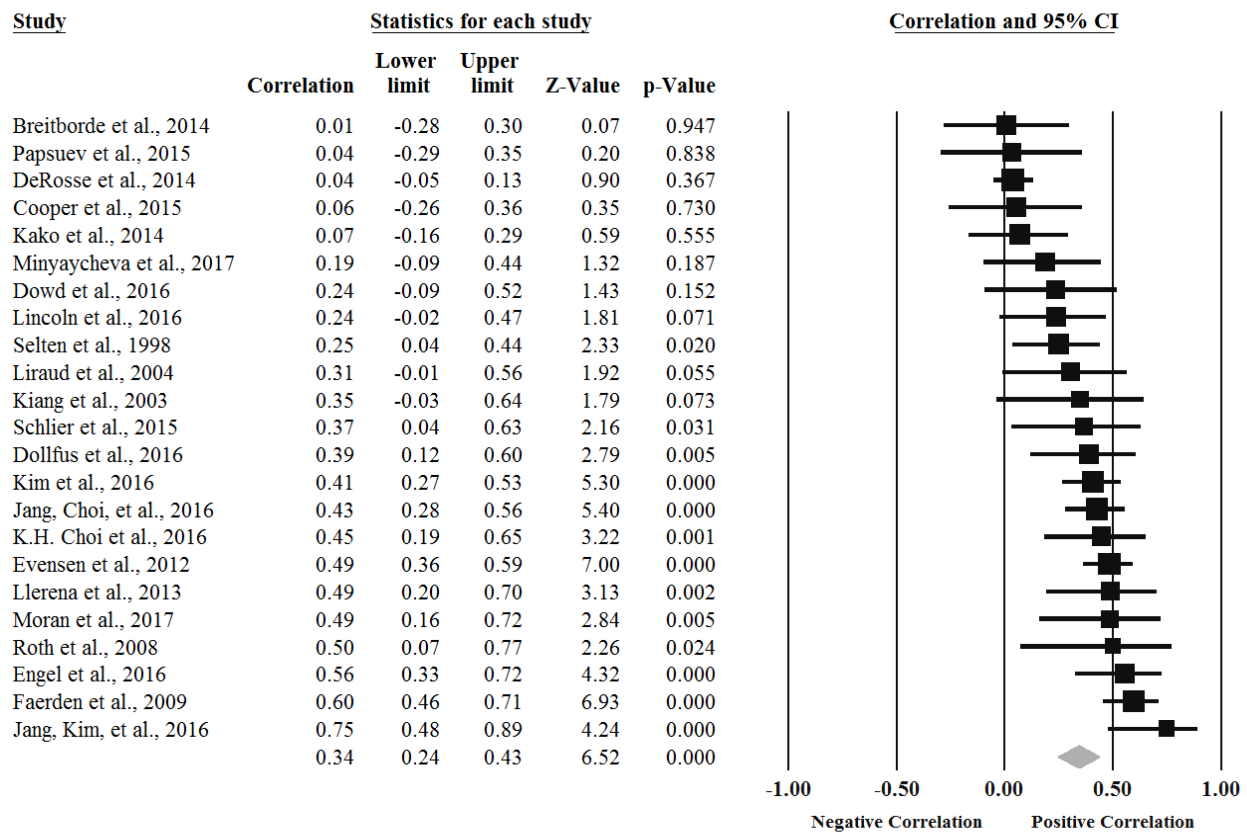


Figure 6. Forest plot of the meta-analysis examining the relationship between self-reported and clinician-rated intrinsic motivation measures ( $k = 4$ )

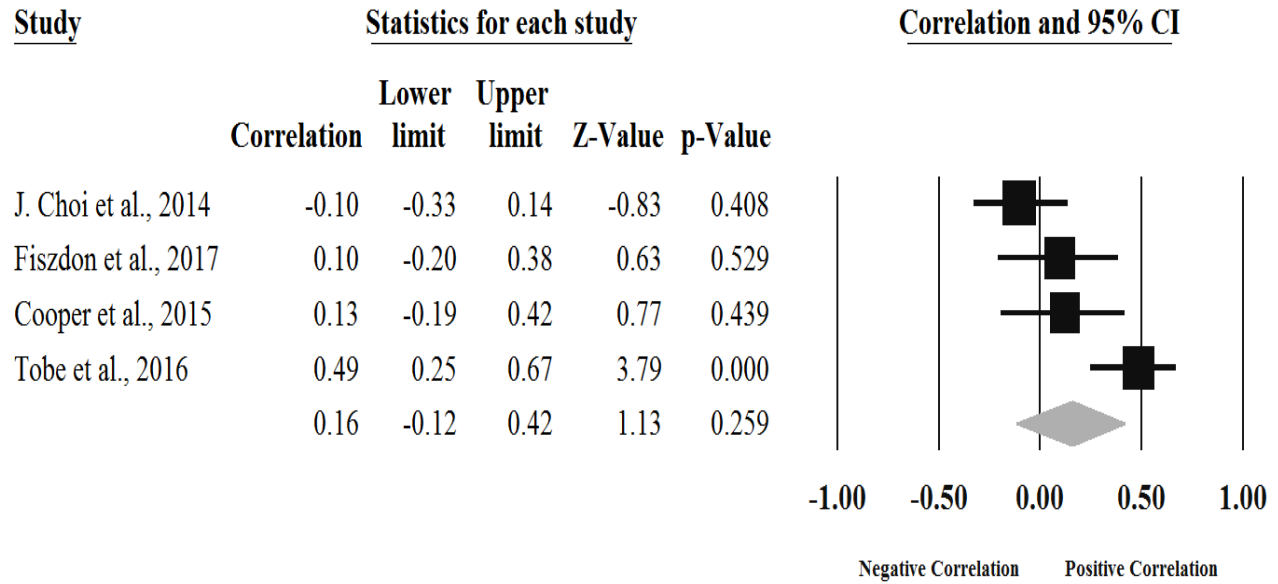


Table 1 *Overall characteristics across included samples (k = 45)*

<b>Sample Characteristics</b>	<b>Mean (SD)</b>	<b>K</b>
Age	38.0 (6.6)	45
Percent Female	37.6 (9.7)	45
Percent Diagnosis <sup>a</sup>		38
Schizophrenia	83.8 (16.8)	-
Schizoaffective	12.9 (13.9)	-
Other Psychosis	3.1 (9.2)	-
Length of illness	13.7 (7.0)	21
Chlorpromazine equivalent doses	482.4 (212.2)	21
<b>Study Characteristics</b>	<b>Mean (SD)</b>	<b>K</b>
Data source (k, %)		45
Published Article <sup>b</sup>	42 (93.3)	-
Conference Abstract	3 (6.7)	-
Median Year (range)	2015 (1998-2017)	45
Mean Sample size (range)	61.8 (11-486)	45
Study Location (k, %)		45
Asia	10 (22.2)	-
Europe	11 (24.4)	-
North America	24 (53.3)	-

<sup>a</sup> All included samples had schizophrenia-spectrum disorder diagnoses.<sup>b</sup> Includes studies that were published as both a conference abstract and article.



Table 2 *Summary of Mean Effect Sizes for the Associations Between Motivation Measures*

Association	<i>k</i>	<i>n</i>	ES - <i>r</i>	95% CI	<i>z</i>	<i>Q</i>	<i>I</i> <sup>2</sup>
Self-reported—Clinician-rated	33	2270	.27	[.19, .35]	6.26***	117.23***	72.70
Performance-based—Clinician-rated	11	445	.21	[.10, .32]	3.77***	11.51	13.09
Self-reported—Performance-based	2	128	-.001	[-.21, .21]	-.01	1.26	20.71
IM Self-reported—IM Clinician-rated	4	209	.16	[-.12, .42]	1.13	11.84**	74.67
EM Self-reported—EM Performance-based <sup>a</sup>	0	-	-	-	-	-	-
AM Self-reported—AM Clinician-rated	23	1847	.34	[.24, .43]	6.52***	97.08***	77.34

Note. *k* = number of eligible studies included in the effect size calculation, if applicable; *n* = total sample combined across eligible studies; ES = weighted and averaged correlation coefficient across eligible studies; 95% CI = 95% confidence interval for the mean effect size; *Z* = z-test value for statistical significance of the mean effect size; *Q* = test for presence of heterogeneity; *I*<sup>2</sup> = indicates the extent of between-study variability; IM = intrinsic motivation; EM = extrinsic motivation; AM = amotivation.

<sup>a</sup>This relationship is listed because based on extant motivation measures, this relationship could be examined; however, no available studies have examined this relationship.

\**p* < .05

\*\**p* < .01

\*\*\**p* < .001

Table 3 Moderator Analyses

Association	Moderator	<i>k</i>	<i>B</i>	<i>SE</i>	95% CI	<i>z</i>	<i>I</i> <sup>2</sup>
Self-Reported— Clinician-rated	Mean age	33	-	.006	[-.018, .006]	-.98	68.70
	Mean illness length	17	-	.007	[-.029, .000] <sup>a</sup>	-	59.26
	Mean CPZ equivalent doses	13	.000	.000	[.000, .001]	.19	50.16
	% female	33	.007	.004	[-.002, .016]	1.56	66.86
	% schizophrenia diagnosis	27	-	.003	[-.006, .004]	-.27	57.23
AM Self- Reported—AM Clinician-rated	Mean age	23	.005	.009	[-.013, .022]	.50	76.50
	Mean illness length	10	-	.016	[-.038, .023]	-.50	66.74
	Mean CPZ equivalent doses	7	.000	.000	[-.001, .000]	-.90	63.06
	% female	23	.007	.005	[-.003, .017]	1.41	70.72
	% schizophrenia diagnosis	17	.000	.003	[-.005, .006]	.15	52.05

Note. *k* = number of studies reporting the moderator and included in the meta-regression; *B* = regression coefficient; *SE* = standard error; 95% CI = 95% confidence interval for regression coefficient; *z* = z-test value for statistical significance of regression coefficient; *I*<sup>2</sup> = indicates the extent of between-study variability that is unexplained after accounting for the moderator; CPZ = chlorpromazine equivalent doses;

AM = Amotivation.

<sup>a</sup> Full CI: [-.0286, -.0003]

\**p* < .05

### Highlights

- An array of motivation measures has been used in schizophrenia research.
- However, the convergent validity of these assessment methods is mixed.
- These meta-analyses summarized the relationships between motivation measure types.
- Results suggest negligent to medium relationships between motivation measure types.
- These measures are assessing a small amount of a shared underlying construct.